The dynamics of metric-affine gravity

Vincenzo Vitagliano[a], Thomas P. Sotiriou[b], Stefano Liberati[a]

[a]*SISSA-International School for Advanced Studies, Via Bonomea 265, 34136 Trieste, Italy and INFN, Sez. di Trieste, Via Valerio 2, 34127 Trieste, Italy*
[b]*Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, UK*

Abstract

Metric-affine theories of gravity provide an interesting alternative to General Relativity: in such an approach, the metric and the affine (not necessarily symmetric) connection are independent quantities. Furthermore, the action should include covariant derivatives of the matter fields, with the covariant derivative naturally defined using the independent connection. As a result, in metric-affine theories a direct coupling involving matter and connection is also present. The role and the dynamics of the connection in such theories is explored. We employ power counting in order to construct the action and search for the minimal requirements it should satisfy for the connection to be dynamical. We find that for the most general action containing lower order invariants of the curvature and the torsion the independent connection does not carry any dynamics. It actually reduces to the role of an auxiliary field and can be completely eliminated algebraically in favour of the metric and the matter field, introducing extra interactions with respect to general relativity. However, we also show that including higher order terms in the action radically changes this picture and excites new degrees of freedom in the connection, making it (or parts of it) dynamical. Constructing actions that constitute exceptions to this rule requires significant fine tuned and/or extra *a priori* constraints on the connection. We also consider $f(\mathcal{R})$ actions as a particular example in order to show that they constitute a distinct class of metric-affine theories with special properties, and as such they cannot be used as representative toy theories to study the properties of metric-affine

---

Email addresses: `vitaglia@sissa.it` (Vincenzo Vitagliano),
`T.Sotiriou@damtp.cam.ac.uk` (Thomas P. Sotiriou), `liberati@sissa.it` (Stefano Liberati)

gravity.

1. Introduction

In general relativity spacetime geometry is fully described by the metric. That is to say, the metric does not only define distances, which is its primary role, but also defines parallel transport, as it is used to construct the Levi-Civita connection. However, in principle this does not have to be the case. The metric and the connection can be independent quantities. In this case one would need field equations that would determine the dynamics of both the metric and the connection.

How can one construct such a theory? In some textbooks, see for example Refs. [1, 2], an independent variation with respect to the metric and the connection of what is formally the Einstein–Hilbert action is considered as an alternative way to arrive to Einstein's equations. This variation is called Palatini variation. Indeed the variation with respect to the connection leads to a non-dynamical equation fixing the latter to be equal to the Levi-Civita connection of the metric, and under this condition the field equations for the metric become Einstein's equations. However, there is a very crucial implicit assumption: that the matter action does not depend on the connection. This is equivalent to assuming that covariant derivatives of matter fields are defined with the Levi-Civita connection of the metric instead of the independent connection. Then the independent connection does not really carry the geometrical meaning described previously, see Refs. [3, 4] for a discussion.

Note that if one allows the connection to enter the matter action, the resulting theory is no longer general relativity [5]. Additionally, one can easily argue that within a metric-affine setting the Einstein–Hilbert form of the action is not necessarily well motivated anyway: under the assumption that the connection is the Christoffel symbol of the metric, the Einstein–Hilbert action is indeed the unique diffeomorphism invariant action which leads to second order field equations (modulo topological terms and total divergences). However, this is not the case if the connection is allowed to be independent and it is not assumed to be symmetric: in this case there are other invariants one should in principle include in the action, even with the same dimensions as the Ricci scalar.

2

The situation gets more complicated once one decides to consider the role of higher order terms. Again, such actions have been studied mostly under the simplifying but geometrically unappealing assumption that the connection does not enter the matter action. Such theories are dubbed Palatini theories of gravity. Particular attention has been payed to $f(\mathcal{R})$ models, i.e. to actions where the Lagrangian is some algebraic function of the Ricci scalar of the independent connection, $\mathcal{R}$. Such actions were introduced and initially studied by Buchdahl [6]. They have recently attracted a lot of interest as possible infrared modifications of general relativity [7] (see Refs. [8, 9, 10] for reviews). However, Palatini $f(\mathcal{R})$ gravity models with infrared corrections with respect to GR have been shown to be non-viable for various reasons: they are in conflict with the standard model of particle physics [11, 12] and they violate solar system tests as their post-Newtonian metric has an algebraic dependence on the matter fields [13, 14]. Singularities have been shown to arise on the surface of well known spherically symmetric matter configurations [15], which render the theory at best incomplete and provide a very strong viability criterion. This criterion is almost independent of the functional form of the Lagrangian, the only exception being Lagrangians with corrections which become important only in the far ultraviolet (as in this case the singularities manifest at scales where non-classical effects take over) [16] . Generalized Palatini theories of gravity have also been considered. For example, Lagrangians of the form $f(\mathcal{R}^{(\mu\nu)}\mathcal{R}_{(\mu\nu)})$ (parentheses indicating symmetrization) were studied in Ref. [17] and Lagrangians of the form $\mathcal{R} + f(\mathcal{R}^{(\mu\nu)}\mathcal{R}_{(\mu\nu)})$ were considered in Ref. [18], with attention being focussed on cosmology. In Refs. [19, 20, 21] Lagrangians of the more general form $f(R, R^{\mu\nu}R_{\mu\nu})$ were studied, see however Ref. [22].

Unlike the exceptional case of the Einstein–Hilbert action as mentioned above, generalized Palatini theories of gravity are distinct from the theories one would get starting from the same action (formally) and applying standard metric variation. One cannot say that their dynamics have been well understood in general. That is because the dynamic of the most well studied class, $f(\mathcal{R})$ are rather peculiar and not representative. Indeed, in Palatini $f(\mathcal{R})$ gravity the independent connection does not carry any dynamics and can be algebraically eliminated in favour of the metric and the matter fields [23, 24]. This result has recently been generalized to $f(R)$ theories with non-symmetric connections, *i.e.* theories that allow for torsion [25]. The lack of extra dynamics with respect to general relativity can also be seen by the fact that Palatini $f(R)$ gravity has been shown to be dynamically equivalent

to Brans–Dicke theory with Brans–Dicke parameter $\omega_0 = -3/2$ [11, 13, 26] (again, irrespectively of how general the connection is allowed to be [25]). This is a particular theory within the Brans–Dicke class in which the scalar does not carry any dynamics and can be algebraically eliminated in favour of the matter fields.

The algebraic elimination of the connection (or the corresponding scalar field in the Brans–Dicke representation) will introduce extra matter interactions [11, 12] and Palatini $f(\mathcal{R})$ theories will essentially be equivalent to general relativity with modified source terms. In fact, this property is what lies in the heart of all the viability issues mentioned earlier [15]. However, this is not a generic property of generalized Palatini gravity, as it has been recently demonstrated in Ref. [22], but just a peculiarity of $f(\mathcal{R})$ actions. Generic higher order actions lead to extra dynamical degrees of freedom.

What we would like to understand here is what happens when one jumps from the Palatini approach, to the more general and better motivated metric-affine approach, where the independent connection is allowed to enter the matter action, define the covariant derivative, and, therefore, retain its geometrical significance. In particular, we would like to understand under which circumstances this connection becomes an auxiliary field, which can be algebraically eliminated, and when it actually does carry dynamics. Note that there are well known examples, such as Einstein–Cartan theory [27] (which is a metric-affine theory with the additional constraint that the connection is metric, but not symmetric), where the independent connection can be eliminated algebraically, leading to general relativity with extra matter interactions. In this specific case, this is a four-fermion interaction. See also Ref. [28] for an example of a more general action with the same property. What happens for more general theories, however, and especially how the dynamics of the connection will be affected by considering higher order terms in the action, has not been systematically understood.

In order to address this issue we follow an approach motivated by effective field theory. We will consider the metric-affine action as an effective action, possibly arising from some more fundamental theory at some appropriate limit. We will then employ power counting in order to construct the most general action order by order. This will allow us to arrive at model independent statements and avoid considering fine-tuned actions, which can lead to misleading results.

The paper is organized in the following way. Section 2 is devoted to presenting our conventions and briefly reviewing the metric-affine set up. In

4

section 3 we construct the most general second order action and show that it does not lead to a dynamical connection. In section 4 we move on to consider actions with higher order invariants and we show that, remarkably, the situation changes radically and degrees of freedom residing in the connection become excited. In section 5 we consider $f(\mathcal{R})$ actions as a particular example, as they have been extensively studied in the literature, even in the metric-affine setting [4]. We do not consider them because we expect them to be representative examples of the dynamics of metric-affine gravity. On the contrary, our intention is to explicitly demonstrate that they are not. Section 6 contains a discussion of our results and our conclusions.

2. General setup for metric-affine theories

We start by clarifying our notation and conventions. The covariant derivative of the connection $\Gamma^{\rho}{}_{\mu\nu}$ acting on a tensor is defined as

$$\nabla_{\mu}A^{\nu}{}_{\sigma} = \partial_{\mu}A^{\nu}{}_{\sigma} + \Gamma^{\nu}{}_{\alpha\mu}A^{\alpha}{}_{\sigma} - \Gamma^{\alpha}{}_{\sigma\mu}A^{\nu}{}_{\alpha}\,. \tag{1}$$

It is important to stress that the position of indices must be taken very carefully into account, since the connection are not assumed to be symmetric. The antisymmetric part of the connection is commonly referred to as the Cartan torsion tensor

$$S_{\mu\nu}{}^{\lambda} \equiv \Gamma^{\lambda}{}_{[\mu\nu]}\,. \tag{2}$$

The failure of the connection to covariantly conserve the metric is measured by the non-metricity tensor

$$Q_{\lambda\mu\nu} \equiv -\nabla_{\lambda}g_{\mu\nu}\,. \tag{3}$$

Using the connection one can construct the Riemann tensor

$$\mathcal{R}^{\mu}{}_{\nu\sigma\lambda} = -\partial_{\lambda}\Gamma^{\mu}{}_{\nu\sigma} + \partial_{\sigma}\Gamma^{\mu}{}_{\nu\lambda} + \Gamma^{\mu}{}_{\alpha\sigma}\Gamma^{\alpha}{}_{\nu\lambda} - \Gamma^{\mu}{}_{\alpha\lambda}\Gamma^{\alpha}{}_{\nu\sigma}\,. \tag{4}$$

which has no dependence on the metric. Notice that the Riemann tensor here has only one obvious symmetry: it is antisymmetric in the last two indices. All other symmetries one might by accustomed to from general relativity are not present for an arbitrary connection [32]. Without any use of the metric we can also define as $\mathcal{R}_{\mu\nu}$ the Ricci tensor built with the connection $\Gamma^{\rho}{}_{\mu\nu}$

$$\mathcal{R}_{\mu\nu} \equiv \mathcal{R}^{\lambda}{}_{\mu\lambda\nu} = \partial_{\lambda}\Gamma^{\lambda}{}_{\mu\nu} - \partial_{\nu}\Gamma^{\lambda}{}_{\mu\lambda} + \Gamma^{\lambda}{}_{\sigma\lambda}\Gamma^{\sigma}{}_{\mu\nu} - \Gamma^{\lambda}{}_{\sigma\nu}\Gamma^{\sigma}{}_{\mu\lambda}\,. \tag{5}$$

5

$\mathcal{R} = g^{\mu\nu}\mathcal{R}_{\mu\nu}$ is the corresponding Ricci scalar.

Note that there is an intrinsic ambiguity in the definition of the Ricci tensor in metric-affine theories as the limited symmetries of the Riemann tensor allow now for an alternative definition as

$$\hat{\mathcal{R}}_{\mu\nu} \equiv \mathcal{R}^{\sigma}{}_{\sigma\mu\nu} = -\partial_{\nu}\Gamma^{\sigma}{}_{\sigma\mu} + \partial_{\mu}\Gamma^{\sigma}{}_{\sigma\nu}. \qquad (6)$$

This tensor is called the homothetic curvature. For a symmetric connection it is equal to the antisymmetric part of $\mathcal{R}_{\mu\nu}$ and, therefore, it need not be separately considered. This is not the case for a non-symmetric connection. Note however, that the homothetic curvature is fully antisymmetric and as such it leads to a vanishing scalar when contracted with the metric.[1]

As already mentioned in the Introduction, the key characteristic of metric-affine gravity is that the affine connection $\Gamma^{\rho}{}_{\mu\nu}$ is not assumed to have any *a priori* relation with the metric. On the other hand, it is assumed to define parallel transport and the covariant derivative of matter fields, so it inevitably enters the matter action, see Ref. [3] for a discussion. That is, in metric-affine gravity couplings between the connection and the matter fields are allowed. This is the main difference from (generalized) Palatini theories of gravity, as mentioned earlier. The action will, therefore be of the following general form

$$S = S_G + S_M = \int d^4x \sqrt{-g} \left[\mathcal{L}_G(g_{\mu\nu}, \Gamma^{\rho}{}_{\mu\nu}) + \mathcal{L}_M \left(g_{\mu\nu}, \Gamma^{\rho}{}_{\mu\nu}, \psi \right) \right], \qquad (7)$$

where g is the determinant of the metric $g_{\mu\nu}$, ψ collectively denotes the matter fields, and S_M is the matter action. We have written the dependence of $\mathcal{L}_M$ on the various fields explicitly to avoid confusion here, but we will suppress it from now on and just use S_M instead in order to lighten the notation. Clearly, specific choices of matter fields will not couple to the connection of course, such as a scalar field or a gauge field. See Ref. [4] for a detailed discussion on such matters and on the general characteristics of metric-affine gravity.

---

[1] See Ref. [4] for a more detailed discussion about the ambiguities in the definition of the Ricci tensor. Note also that, unlike the usual Ricci tensor, the homothetic curvature tensor has a direct physical interpretation: it measures the change of the length of a vector when it is transported along a closed loop. When the homothetic curvature vanishes, the connection is volume preserving, *i.e.* lengths and volumes do not change during parallel transport. We thank Yuri Obukhov for bringing this to our attention.

One now needs to specify the exact form of the Lagrangian $\mathcal{L}_G$. In Ref. [5] an action linear in $\mathcal{R}$ was consider and in Ref. [4] the most general $f(\mathcal{R})$ family was studied extensively. Instead of an *ad hoc* choice inspired by some similarity with the Einstein–Hilbert action and its generalizations, we would like to follow here an effective field theory approach in order consider the most general action possible at each order. To construct this action, we should carry on a power counting analysis which will reveal the whole set of appropriate terms order by order. We set $c = 1$ and we can choose the engineering dimensions

$$[dx] = [dt] = [l] \tag{8}$$

where l is a place holder symbols with dimension of a length. Then we have

$$[g_{\mu\nu}] = [1]\,, \quad [\sqrt{-g}dx^4] = [l^4]\,, \quad [\Gamma^\lambda{}_{\mu\nu}] = [l^{-1}]\,, \quad [\mathcal{R}_{\mu\nu}] = [l^{-2}]\,. \tag{9}$$

Now consider as a simple example the action

$$S_G = \frac{1}{l_p^2} \int dx^3 dt \sqrt{-g}\mathcal{R}\,. \tag{10}$$

Requiring that this action is dimensionless implies that the coupling constant l_p must have dimensions of a length and can be naturally associated to the Planck length. What we mean by order of the gravitational theory is also clear now: we mean the highest order in l_p^{-1} powers appearing in the Lagrangian (which, since one cannot choose the metric and the connection to be dimensionless at the same time, does not correspond to the order in its derivatives).

3. Second order action

Clearly the action written above is not the most general one we could write in metric-affine gravity. It is just an example inspired by the analogy with standard GR and the Einstein–Hilbert action. To begin with, we could include a cosmological constant term, which is of lower order. But such a term would not play any important role in our arguments so we will omit it for simplicity. What other terms can we write at the second order? Under the assumption that the connection is torsionless and metric compatible (Levi-Civita), there exist no other term which respects diffeomorphism invariance, as it is well known. But, in the more general metric-affine setting we are considering here, there is at least two more tensors one could imagine using in order to construct invariants:

- The aforementioned "second Ricci" tensor $\hat{\mathcal{R}}_{\mu\nu}$. However, this tensor has dimensions $[l^{-2}]$ and is antisymmetric, so there is no quantity one can construct out of it at second order;

- the Cartan torsion tensor of eq. (2), which has the same dimensions as $\Gamma^\lambda_{\ \mu\nu}$. Therefore, terms with one derivative of $S_{\mu\nu}^{\ \ \lambda}$ or terms quadratic in $S_{\mu\nu}^{\ \ \lambda}$ will be of the same order as $\mathcal{R}$.

Due to the symmetries of $S_{\mu\nu}^{\ \ \lambda}$ there is only a single term with a derivative we can write

$$g^{\mu\nu}\nabla_\mu S_{\nu\sigma}^{\ \ \sigma} \, . \tag{11}$$

For the same reason, there are just three terms quadratic in $S_{\mu\nu}^{\ \ \lambda}$ one can write

$$g^{\mu\nu} S_{\mu\lambda}^{\ \ \lambda} S_{\nu\sigma}^{\ \ \sigma} \, , \qquad g^{\mu\nu} S_{\mu\lambda}^{\ \ \sigma} S_{\nu\sigma}^{\ \ \lambda} \, , \qquad g^{\mu\alpha} g^{\nu\beta} g_{\lambda\gamma} S_{\mu\nu}^{\ \ \lambda} S_{\alpha\beta}^{\ \ \gamma} \, . \tag{12}$$

Note that the term in eq. (11) has been considered by Papapetrou and Stachel in [30].

A subtle point is the following. The term in eq. (11) is not a total divergence as ∇_μ is not defined with the Levi-Civita connection of the metric. On the other hand, one can think to decompose the connection as

$$\Gamma^\lambda_{\ \mu\nu} = \left\{ ^\lambda_{\ \mu\nu} \right\} + C^\lambda_{\ \mu\nu} \, , \tag{13}$$

i.e. in its Levi-Civita part and the rest. Now, using this decomposition we can split the covariant derivative in (11) in a metric compatible part, which will lead to a total divergence, and the rest, which will lead to terms consisting of contractions between $C^\lambda_{\ \mu\nu}$ and the Cartan torsion tensor. Since the non-metricity is not zero, these terms are different than the ones already considered above in eq. (12). Thus the term in eq. (11) is non-trivial.

This brings us to another puzzle though: $C^\lambda_{\ \mu\nu}$ is a tensor, so why not consider terms constructed with it as well? Actually, $C^\lambda_{\ \mu\nu}$ can always be decomposed in terms of torsion $S_{\mu\nu}^{\ \ \lambda}$ and non-metricity $Q_{\lambda\mu\nu}$, so the question then reduces to whether we should also consider terms constructed with $Q_{\lambda\mu\nu}$ or not. From a power counting/field theory perspective nothing prevents us from doing so, and these would indeed be terms of the same order. However, from this perspective we should also consider, for instance, the Ricci tensor of the metric $R_{\mu\nu}$. In fact, $Q_{\lambda\mu\nu}$ and $R_{\mu\nu}$ share a common characteristic which is crucial for our discussion: They cannot be expressed without using derivatives of the metric (even if instead of (3) one tries to define $Q_{\lambda\mu\nu}$ using

8

the connection, then still the Levi-Civita connection would be needed as well). Therefore, the puzzle reduces to whether or not we should be considering invariants constructed with derivatives of the metric.

Clearly, field theoretic considerations cannot give an answer to this question. Such terms should be considered unless we are willing to invoke some principle excluding them, along the line of minimal coupling in general relativity. Such a principle has been discussed in Ref. [4]. In simple terms it would be the requirement that the metric be used only for raising and lowering indices. We choose to follow this prescription here, as it seems sensible from a geometrical perspective (the purpose of the metric being to measure distances) and it significantly reduces the number of terms one can consider.

Another way to reduce the number of terms without invoking a minimal coupling principle would be to require that the connection be metric compatible. This would force $Q_{\lambda\mu\nu}$ to vanish, without necessarily implying torsion has to vanish as well. We would then remain with exactly the same terms written above. However, in this case the term in eq. (11) would indeed differ from the first term in eq. (12) only by a total surface term and one would be able to omit it.

Let us consider the consider the most general second-order action we have just constructed in our setting

$$
S = \frac{1}{16\,\pi\,l_p^2} \int dx^4 \sqrt{-g} \quad \left(g^{\mu\nu}\mathcal{R}_{\mu\nu} + a_1 g^{\mu\nu}\nabla_\mu S_{\nu\sigma}{}^\sigma + a_2 g^{\mu\nu} S_{\mu\lambda}{}^\lambda S_{\nu\sigma}{}^\sigma \right. \tag{14}
$$

$$
\left. + a_3 g^{\mu\nu} S_{\mu\lambda}{}^\sigma S_{\nu\sigma}{}^\lambda + a_4 g^{\mu\alpha}g^{\nu\beta}g_{\lambda\gamma} S_{\mu\nu}{}^\lambda S_{\alpha\beta}{}^\gamma \right) + S_M \,,
$$

where the a_i's represent the various coupling constant. Varying independently with respect to metric and connection yields

$$
\mathcal{R}_{(\mu\nu)} - \frac{1}{2}g_{\mu\nu}\mathcal{R} + a_1\left\{\nabla_{(\mu}S_{\nu)} - \frac{1}{2}g_{\mu\nu}g^{\alpha\beta}\nabla_\alpha S_\beta\right\}
$$

$$
+ a_2\left\{-\frac{1}{2}g_{\mu\nu}S_\alpha S^\alpha + S_\mu S_\nu\right\} + a_3\left\{-\frac{1}{2}g_{\mu\nu}g^{\alpha\beta}S_{\alpha\lambda}{}^\sigma S_{\beta\sigma}{}^\lambda + S_{\mu\lambda}{}^\sigma S_{\nu\sigma}{}^\lambda\right\}
$$

$$
+ a_4\left\{-\frac{1}{2}g_{\mu\nu}S_{\rho\sigma\lambda}S^{\rho\sigma\lambda} + 2S_{\alpha\mu}{}^\lambda S^\alpha{}_{\nu\lambda} - S_{\rho\sigma\mu}S^{\rho\sigma}{}_\nu\right\} = \kappa T_{\mu\nu}\,, \tag{15}
$$

$$
\frac{1}{\sqrt{-g}}\left[-\nabla_\lambda\left(\sqrt{-g}g^{\mu\nu}\right) + \nabla_\sigma\left(\sqrt{-g}g^{\sigma\mu}\right)\delta^\nu{}_\lambda - a_1\nabla_\alpha\left(\sqrt{-g}g^{\alpha[\mu}\right)\delta^{\nu]}{}_\lambda\right]
$$

$$
+ (2-a_1)\,g^{\mu\nu}S_\lambda - 2S^{(\mu}\delta^{\nu)}{}_\lambda + 2(a_1+a_2-1)S^{[\mu}\delta^{\nu]}{}_\lambda + 2a_3 g^{\alpha[\mu}S_{\alpha\lambda}{}^{\nu]}
$$

$$
+ 2a_4 g^{\alpha[\mu}g^{\nu]\beta}g_{\lambda\gamma}S_{\alpha\beta}{}^\gamma = \kappa\Delta_\lambda{}^{\mu\nu}\,. \tag{16}
$$

where $S_\alpha \equiv S_{\alpha\beta}{}^\beta$, $\kappa = 8\pi\, l_p^2$ and

$$T_{\mu\nu} \equiv -\frac{2}{\sqrt{-g}}\frac{\delta S_M}{\delta g^{\mu\nu}}\,, \qquad \Delta_\lambda{}^{\mu\nu} \equiv -\frac{2}{\sqrt{-g}}\frac{\delta S_M}{\delta \Gamma^\lambda{}_{\mu\nu}}\,. \qquad (17)$$

$\Delta_\lambda{}^{\mu\nu}$ is now as the *hypermomentum* and, as already identified in [31], it encapsulates all the information related to the spin angular momentum of matter, the intrinsic part of dilation current and the shear current. $T_{\mu\nu}$ on the other hand is sometimes referred to as the stress-energy tensor, in analogy with general relativity. However, it should be stressed that this terminology might be misleading within the metric-affine framework as this tensor does not have the properties usually associated with the stress-energy tensor is general relativity. For instance, it is not necessarily divergence free, it does not reduce to the special relativistic stress energy tensor at a suitable limit and of course it does describe but some properties of matter, given the existence of $\Delta_\lambda{}^{\mu\nu}$ as well. In fact, it is best if it is just considered as nothing more than a short hand notation for the functional derivative of the matter action with respect to the metric.

Our present aim is to check whether it is possible to fully eliminate the connection from the field equations. Let us consider the contraction of λ index in (16) respectively with μ and ν

$$\frac{3}{2}\frac{a_1}{\sqrt{-g}}\nabla_\mu\left(\sqrt{-g}g^{\mu\nu}\right) = \kappa\Delta_\mu{}^{\mu\nu} + (4a_1 + 3a_2 + a_3 + 2a_4)S^\nu\,, \qquad (18)$$

$$\frac{6 - 3a_1}{2\sqrt{-g}}\nabla_\nu\left(\sqrt{-g}g^{\mu\nu}\right) = \kappa\Delta_\nu{}^{\mu\nu} - (2a_1 + 3a_2 + a_3 + 2a_4 - 6)S^\mu\,. \qquad (19)$$

Combining these two equations gives S_ν and the trace $\nabla_\mu\left(\sqrt{-g}g^{\mu\nu}\right)$ as functions of the hypermomentum

$$S^\nu = \frac{\kappa}{(1 - a_1)a_1 + 3a_2 + a_3 + 2a_4}\left[(a_1 - 1)\Delta^\nu - \tilde{\Delta}^\nu\right]\,, \qquad (20)$$

$$\frac{1}{\sqrt{-g}}\nabla_\mu\left(\sqrt{-g}g^{\mu\nu}\right) = \frac{2}{3}\left\{\kappa\Delta^\nu + \frac{\kappa(a_1 + 3)\left[(a_1 - 1)\Delta^\nu - \tilde{\Delta}^\nu\right]}{(1 - a_1)a_1 + 3a_2 + a_3 + 2a_4}\right\}\,, (21)$$

where we defined the two quantities $\Delta^\mu \equiv \Delta_\alpha{}^{(\alpha\mu)}$ and $\tilde\Delta^\mu \equiv \Delta_\alpha{}^{[\alpha\mu]}$. Eq. (21) can be inserted in (16) to eliminate the second and the third term in order to get

$$\frac{1}{\sqrt{-g}}\left[-\nabla_\lambda\left(\sqrt{-g}\,g^{\mu\nu}\right)\right] + 2a_3 g^{\alpha[\mu}S_{\alpha\lambda}{}^{\nu]} + 2a_4 g^{\alpha[\mu}g^{\nu]\beta}g_{\lambda\gamma}S_{\alpha\beta}{}^\gamma =$$

$$= \kappa\Delta_\lambda{}^{\mu\nu} - \frac{2}{3}\left[\kappa\Delta^\mu + (a_1+3)S^\mu\right]\delta^\nu{}_\lambda + \frac{2}{3}a_1\left\{\kappa\Delta^{[\mu} + (a_1+3)S^{[\mu}\right\}\delta^{\nu]}{}_\lambda -$$

$$- (2-a_1)\,g^{\mu\nu}S_\lambda + 2S^{(\mu}\delta^{\nu)}{}_\lambda - 2(a_1+a_2-1)S^{[\mu}\delta^{\nu]}{}_\lambda\,, \qquad (22)$$

while we will refrain from replacing S_ν for compactness.

Using the identities

$$\nabla_\mu\sqrt{-g} = \partial_\mu\sqrt{-g} - \Gamma^\alpha{}_{\alpha\mu}\sqrt{-g} \qquad (23)$$

and

$$g_{\mu\nu}\partial_\lambda(\sqrt{-g}\,g^{\mu\nu}) = 2\sqrt{-g}\,\partial_\lambda\ln\sqrt{-g} \qquad (24)$$

we can write the trace of eq. (22) with the metric in the μ and ν indices as

$$\frac{\partial_\lambda\sqrt{-g}}{\sqrt{-g}} = -\frac{1}{2}\kappa g_{\mu\nu}\Delta_\lambda{}^{\mu\nu} + \frac{1}{3}\kappa g_{\mu\lambda}\Delta^\mu + \left(4 - \frac{5}{3}a_1\right)S_\lambda + \Gamma^\alpha{}_{\alpha\lambda}\,. \qquad (25)$$

Eliminating the density related terms and suitably lowering the indices eq. (22) can eventually take the form

$$\partial_\lambda g_{\sigma\rho} - \Gamma^\mu{}_{\rho\lambda}g_{\mu\sigma} - \Gamma^\nu{}_{\sigma\lambda}g_{\nu\rho} + 2a_3 S_{[\sigma|\lambda|\rho]} + 2a_4 S_{\sigma\rho\lambda} =$$

$$= g_{\sigma\rho}\left(4 - \frac{5}{3}a_1\right)S_\lambda + \frac{1}{3}\kappa g_{\sigma\rho}\Delta_\lambda - \frac{1}{2}\kappa g_{\sigma\rho}\Delta_\lambda{}^\mu{}_\mu +$$

$$+ \kappa\Delta_{\lambda\sigma\rho} - (2-a_1)\,g_{\sigma\rho}S_\lambda + 2S_{(\sigma}g_{\rho)\lambda} - 2(a_1+a_2-1)S_{[\sigma}g_{\rho]\lambda} +$$

$$+ \frac{2}{3}(a_1+3)\left(a_1 S_{[\sigma}g_{\rho]\lambda} - S_\sigma g_{\rho\lambda}\right) + \frac{2}{3}\kappa\left(a_1\Delta_{[\sigma}g_{\rho]\lambda} - \Delta_\sigma g_{\rho\lambda}\right)\,. \qquad (26)$$

We can now split this last expression in its antisymmetric and symmetric part with respect to the two indices σ and ρ

$$2a_3 S_{[\sigma|\lambda|\rho]} + 2a_4 S_{\sigma\rho\lambda} = \Theta_{\lambda\sigma\rho}\,,$$

$$\partial_\lambda g_{\sigma\rho} - \Gamma^\mu{}_{\rho\lambda}g_{\mu\sigma} - \Gamma^\nu{}_{\sigma\lambda}g_{\nu\rho} = \kappa\Delta_{\lambda(\sigma\rho)} - \frac{2}{3}\left[\kappa\Delta_{(\sigma} + a_1 S_{(\sigma}\right]g_{\rho)\lambda}$$

$$- g_{\sigma\rho}\left[\left(\frac{2}{3}a_1 - 2\right)S_\lambda + \frac{1}{2}\kappa\Delta_\lambda{}^\mu{}_\mu - \frac{1}{3}\kappa\Delta_\lambda\right]\,, \qquad (27)$$

11

where we have introduced the short hand notation

$$\Theta_{\lambda\sigma\rho} \equiv \kappa\Delta_{\lambda[\sigma\rho]} + \frac{2}{3}(a_1 - 1)\left[\kappa\Delta_{[\sigma} + \left(a_1 - \frac{3a_2}{a_1 - 1}\right)S_{[\sigma}\right]g_{\rho]\lambda}. \tag{28}$$

Adding suitable permutations of (27) and (27) we obtain

$$S_{\rho\nu\mu} = \frac{a_3}{2a_3(a_3 + a_4) - 4a_4^2}\left[\Theta_{\nu\rho\mu} - \Theta_{\rho\nu\mu} - \left(2\frac{a_4}{a_3} - 1\right)\Theta_{\mu\rho\nu}\right], \tag{29}$$

$$\Gamma^{\xi}_{(\sigma\rho)} = \left\{^{\xi}_{\sigma\rho}\right\} + 2S_{(\rho}{}^{\xi}{}_{\sigma)} - \frac{1}{2}\kappa g^{\xi\lambda}(-\Delta_{\lambda(\sigma\rho)} + \Delta_{\rho(\sigma\lambda)} + \Delta_{\sigma(\rho\lambda)})$$

$$-\frac{\kappa}{3}g^{\xi\lambda}\left(2\Delta_{(\sigma}g_{\rho)\lambda} - 3\Delta_{\lambda}g_{\sigma\rho}\right) + g^{\xi\lambda}\left[S_{\lambda}g_{\sigma\rho} + \left(\frac{2}{3}a_1 - 2\right)S_{(\sigma}g_{\rho)\lambda}\right]$$

$$-\frac{\kappa}{4}g^{\xi\lambda}\left[g_{\sigma\rho}\Delta_{\lambda\mu}{}^{\mu} - 2g_{\lambda(\rho}\Delta_{\sigma)\mu}{}^{\mu}\right]. \tag{30}$$

Eqs. (29) and (30) give the antisymmetric and symmetric parts of the connection algebraically in terms of the hypermomentum and the metric. Under the condition that the matter action depends at most linearly on the connection, the above statement is equivalent to saying that we have algebraically expressed the connection in terms of the matter fields and the metric. This assumption is indeed satisfied for all common matter actions, such as scalar and gauge field, in which the matter action does not depend on the connection, and fermions, where the matter action is linear in the connection. This condition can be violated for some fields, for example massive vector fields, especially if not trivial couplings between the connection and the matter are introduced. However, as long as the matter action contains only first order derivatives of the matter fields (in order for the matter fields to satisfy second order equations of motion), $\Delta_{\lambda}{}^{\mu\nu}$ will only depend algebraically on the connection. This implies that, even though some more complicated manipulations will be required, the connection can always be expressed algebraically in terms of the matter field and the metric (at least at the component level).

This establishes that the independent connection in (up to) second order metric-affine actions does not carry any dynamics and it can be algebraically eliminated. Consider now using eqs. (30) and (29) to completely eliminate the connection in eq. (15). One would then get an equations of the form

$$R_{\mu\nu} - \frac{1}{2}R\,g_{\mu\nu} = \kappa\mathcal{T}_{\mu\nu}, \tag{31}$$

where $R_{\mu\nu}$ and R are the Ricci tensor and the Ricci scalar of the metric $g_{\mu\nu}$ respectively, and $\mathcal{T}_{\mu\nu}$ will be some a second rank tensor which depends on the

12

metric, $\Delta_\lambda{}^{\mu\nu}$ and $T_{\mu\nu}$. The expression for $\mathcal{T}_{\mu\nu}$ in terms of these three quantities is rather lengthy and we will refrain from writing it here. However, it should already be clear that the theory described by eq. (31) is general relativity with modified matter interactions. For fields for which the hypermomentum vanishes, $\mathcal{T}_{\mu\nu} = T_{\mu\nu}$.

4. Higher orders

We can now move on to higher orders. Since the connection has three indices and the derivative one index, there is no $[l^{-3}]$ scalar quantity one can construct out of them. Similarly, one cannot construct an $[l^{-3}]$ scalar quantity using curvature invariants. The next order is $[l^{-4}]$. The terms that could straightforwardly lead to invariants after (several) contractions with the metric are

$$\mathcal{R}^\alpha{}_{\beta\gamma\delta}\mathcal{R}^\mu{}_{\nu\lambda\sigma}, \qquad \nabla_\mu\nabla_\nu\mathcal{R}^\alpha{}_{\beta\gamma\delta}, \qquad \mathcal{R}^\alpha{}_{\beta\gamma\delta}S_{\mu\nu}{}^\lambda S_{\tau\omega}{}^\rho, \qquad \mathcal{R}^\alpha{}_{\beta\gamma\delta}\nabla_\rho S_{\mu\nu}{}^\lambda$$
$$S_{\mu\nu}{}^\lambda\nabla_\rho\mathcal{R}^\alpha{}_{\beta\gamma\delta}, \qquad S_{\mu\nu}{}^\lambda S_{\alpha\beta}{}^\sigma S_{\gamma\delta}{}^\kappa S_{\tau\omega}{}^\rho, \qquad S_{\mu\nu}{}^\lambda S_{\alpha\beta}{}^\sigma\nabla_\rho S_{\gamma\delta}{}^\kappa$$
$$S_{\mu\nu}{}^\lambda\nabla_\rho\nabla_\kappa S_{\alpha\beta}{}^\sigma, \qquad \nabla_\rho S_{\mu\nu}{}^\lambda\nabla_\kappa S_{\alpha\beta}{}^\sigma, \qquad \nabla_\mu\nabla_\nu\nabla_\rho S_{\alpha\beta}{}^\sigma, \qquad (32)$$

Clearly each of these terms can lead to various invariants. It goes beyond the purpose of the paper to list all possible terms.[2] However, before going further, the following subtle points are worth mentioning:

1. Due to the symmetries (or lack thereof) of the Riemann tensor when constructed with an independent connection, there are more invariants than in the purely metric case. For example $\mathcal{R}_{\mu\nu}$ is not symmetric and hence $\mathcal{R}_{\mu\nu}\mathcal{R}_{\kappa\lambda}g^{\mu\lambda}g^{\nu\kappa}$ and $\mathcal{R}_{\mu\nu}\mathcal{R}_{\kappa\lambda}g^{\mu\kappa}g^{\nu\lambda}$ are not equal.
2. ∇_μ is constructed with the independent connection and, hence, total divergences such as $\nabla_\mu u^\mu$ do not lead to pure surface terms and cannot be discarded.
3. Since the metric is not covariantly conserved by the independent connection taking the covariant derivatives first and contracting, or contracting first and then taking a derivative does not lead to the same result. For example the terms $g^{\mu\nu}g^{\alpha\beta}\nabla_\mu\nabla_\nu\mathcal{R}_{\alpha\beta}$ and $g^{\mu\nu}\nabla_\mu\nabla_\nu\mathcal{R}$ differ.

---

[2]An exhaustive list of all possible second and fourth order invariants one can construct in the more limiting case where the non-metricity vanishes can be found in [29]. Given that our minimal coupling assumption prevents us from the using of the non-metricity to construct invariants (see also below), this list should cover our case as well.

Regarding point (ii) one could propose to split the covariant derivative into a metric covariant derivative, which is a surface term, and the rest, such as in (13). However, writing the rest explicitly would require the use of metric derivatives through the use of the Levi-Civita connection, as discussed above. Something similar can be said about point (iii). The two terms given as an example differ by a term including a covariant derivative of the metric. This raises the question of whether both of them should be considered. As mentioned earlier, whether terms including derivatives of the metric should be included is really a matter of choice that can be answered only by invoking some minimal coupling principle. If one wants to use the metric purely for contracting indices as suggested previously, then the terms including derivatives of the metric should be suppressed.

Let us now move on to consider the effect of the higher order terms on the dynamics of the connection. Considering the most general fourth order action is formidable due to the vast number of invariant one would have to include. However, carefully considering isolated terms of different type can still reveal the complete picture.

Clearly there are term in eq. (32) that would not introduce new degrees of freedom if they were added to action (14) as they do not contain extra derivatives, such as the S^4 term (indices suppressed). Such terms exist at all even orders, e.g. S^{2n} (again indices suppressed). On the other hand $[l^{-4}]$ terms which contain two derivatives of the Cartan torsion tensor, such as $(\nabla S)^2$ (indices suppressed) would inevitable make the torsion dynamical.

What about fourth order curvature invariants? Let us for the moment set aside the term $\mathcal{R}^2$, since it belongs to the general $f(\mathcal{R})$ class, which we will discuss extensively later, and as we will see it constitutes a rather special case. A much more characteristic example to consider, which is simple enough to keep calculations tractable and yet general enough to give us the bigger picture is the following

$$S = \frac{1}{16\,\pi\,l_p^2} \int dx^4 \sqrt{-g} \left[\mathcal{R} + l_p^2 \mathcal{R}_{\mu\nu} \mathcal{R}_{\kappa\lambda} (ag^{\mu\kappa}g^{\nu\lambda} + bg^{\mu\lambda}g^{\nu\kappa}) \right] + S_M \quad (33)$$

As mentioned earlier, when $\mathcal{R}_{\mu\nu}$ is not symmetric, as in our case, the 2 terms in the parenthesis will not lead to the same invariant. In fact, the action can be re-written as

$$S = \frac{1}{16\,\pi\,l_p^2} \int dx^4 \sqrt{-g} \left[\mathcal{R} + l_p^2 c_1 \mathcal{R}_{(\mu\nu)} \mathcal{R}^{(\mu\nu)} + l_p^2 c_2 \mathcal{R}_{[\mu\nu]} \mathcal{R}^{[\mu\nu]} \right] + S_M \quad (34)$$

14

where $c_1 = a + b$ and $c_2 = a - b$. This latter form of the action makes the variation easier. The field equations for the metric and the connection are respectively

$$\mathcal{R}_{(\mu\nu)} - \frac{1}{2}\left(\mathcal{R} + l_p^2 c_1 \mathcal{R}_{(\alpha\beta)}\mathcal{R}^{(\alpha\beta)} + l_p^2 c_2 \mathcal{R}_{[\alpha\beta]}\mathcal{R}^{[\alpha\beta]}\right) g_{\mu\nu}$$
$$+ 2l_p^2 c_1 R_{(\alpha\mu)}R_{(\beta\nu)}g^{\alpha\beta} + 2l_p^2 c_2 R_{[\alpha\mu]}R_{[\beta\nu]}g^{\alpha\beta} = \kappa T_{\mu\nu}\,, \qquad (35)$$

$$\frac{1}{\sqrt{-g}}\Big\{ -\nabla_\lambda\left[\sqrt{-g}g^{\mu\nu} + 2\sqrt{-g}\left(l_p^2 c_1 \mathcal{R}^{(\mu\nu)} + l_p^2 c_2 \mathcal{R}^{[\mu\nu]}\right)\right] +$$
$$+ \nabla_\sigma\left[\sqrt{-g}g^{\mu\sigma} + 2\sqrt{-g}\left(l_p^2 c_1 \mathcal{R}^{(\mu\sigma)} + l_p^2 c_2 \mathcal{R}^{[\mu\sigma]}\right)\right]\delta^\nu{}_\lambda\Big\} +$$
$$+ 2S_{\lambda\sigma}{}^\sigma\left[g^{\mu\nu} + 2\left(l_p^2 c_1 \mathcal{R}^{(\mu\nu)} + l_p^2 c_2 \mathcal{R}^{[\mu\nu]}\right)\right]$$
$$- 2S_{\alpha\sigma}{}^\sigma\delta^\nu{}_\lambda\left[g^{\mu\alpha} + 2\left(l_p^2 c_1 \mathcal{R}^{(\mu\alpha)} + l_p^2 c_2 \mathcal{R}^{[\mu\alpha]}\right)\right] +$$
$$+ 4\left(l_p^2 c_1 \mathcal{R}^{(\mu\alpha)} + l_p^2 c_2 \mathcal{R}^{[\mu\alpha]}\right)S_{\alpha\lambda}{}^\nu = \kappa\Delta_\lambda{}^{\mu\nu}\,. \qquad (36)$$

In the previous section we were able to use the field equation for the connection in order to algebraically express the latter in terms of the metric and the matter fields. Inspecting eq. (36), however, one easily realized that, unlike eq. (16), it appears to include derivatives of the connection due to the presence of $\mathcal{R}_{\mu\nu}$. One could think to use eq. (35) in order to algebraically express $\mathcal{R}_{\mu\nu}$ (at least at component level) in terms of the metric and the matter fields (this idea is actually inspired by the specific case of $f(R)$ actions in the more restricted setting of the Palatini formalism where the connection does not couple to the matter — this will be discussed below). If this were the case, one could eliminate $\mathcal{R}_{\mu\nu}$ from eq. (36) and turn it again into an algebraic equation for the connection.

However, this is not possible for generic values of c_1 and c_2, or a and b for the following simple reasons:

- $\mathcal{R}_{\mu\nu}$ is not necessarily symmetric and, therefore, has 16 independent components, whereas eq. (35) leads to only 10 components equation as it is symmetric in μ and ν. Therefore, it cannot be used to determine $\mathcal{R}_{\mu\nu}$ fully, in terms of the metric and the components of $T_{\mu\nu}$.

- $T_{\mu\nu}$ is not necessarily independent of the connection, as it may include covariant derivatives of certain matter fields. Therefore, even if one would impose such conditions so that eq. (35) could be solved algebraically to give $\mathcal{R}_{\mu\nu}$ in term of the metric and $T_{\mu\nu}$, e.g. impose the

constraint $\mathcal{R}_{[\mu\nu]} = 0$ *a priori*, that would not actually help in algebraically expressing the connection as a function of the matter fields and the metric (at least for generic matter fields).

It should then be clear that the independent connection cannot be eliminated in metric-affine gravity once generic higher order curvature invariants have been added.

The same issue has been considered in Ref. [22] for the simpler case of generalized Palatini gravity, i.e. under the assumption that connection does not enter the matter action. This would corresponds to a vanishing $\Delta_\lambda{}^{\mu\nu}$. The first of the difficulties just discussed is still present in this case when trying to eliminate the connection algebraically by the procedure described above. However, since $T_{\mu\nu}$ is independent of the connection in generalized Palatini gravity, the second difficulty raised here is not really an issue. Hence, it is easier in this framework to write down exceptional Lagrangians for which the connection can be eliminated (it is just an auxiliary field). We refer the reader to Ref. [22] for more details. We refrain here from discussing similar exceptions or special cases for metric-affine gravity, as this would require severe fine tuning and/or *a priori* constraints.

Also, we shall not consider explicitly the effect of the mixed terms which include both the Cartan torsion tensor and the Riemann or the Ricci tensor, as this would not add anything new to the qualitative understanding we presented so far. What should be clear by now is that the presence of terms including derivatives of the Cartan torsion tensor or higher order curvature invariants generically leads to a dynamical connection. Therefore, higher than second order actions generically lead to more dynamical degrees of freedom.

5. Metric-affine $f(\mathcal{R})$ gravity as a special case

Metric-affine $f(\mathcal{R})$ theories of gravity have been extensively studied lately [4]. They constitute a distinct class within higher order actions, in the sense that they allows one to treat terms of different and arbitrarily high order on the same footing. Therefore, even though within the metric-affine setup there is no reason to single out $f(\mathcal{R})$ actions as better motivated ones — on the contrary, restricting an action to be of this type requires fine tuning — their simplicity is indeed a good argument for adopting them as toy-models from which to extract general lessons. On the other hand, exactly because

16

they are so special, it is dubious whether $f(\mathcal{R})$ actions can be considered as representative higher order metric-affine theories from the point of view of their dynamics. This is something that is worth exploring further, which is part of our motivation for considering them separately here.

The other part of our motivation comes from the observation that in the simpler setting of generalized Palatini gravity, where the connection does not enter the matter action, the whole $f(\mathcal{R})$ class constitutes an exception for which the independent connection does not carry dynamics and can be algebraically eliminated [3, 25]. This is true even if the connection is not assumed to be symmetric [25]. It is, hence, worth exploring in detail what happens in the more general metric-affine framework, in order to avoid confusion and misconceptions.

The action for $f(\mathcal{R})$ theories reads

$$S = \frac{1}{16\,\pi\,l_p^2} \int d^4x \sqrt{-g} f(\mathcal{R}) + S_M \tag{37}$$

This action as it stands cannot lead to consistent field equations in the presence of matter, as the gravity part of the action has a symmetry that is not shared by the matter action. The Ricci scalar of the connection $\mathcal{R}$ remain invariant under the projective transformation

$$\Gamma^{\rho}{}_{\mu\nu} \to \Gamma^{\rho}{}_{\mu\nu} + \delta^{\rho}{}_{\mu}\xi_{\nu} \tag{38}$$

(ξ_{μ} being an arbitrary covariant vector field). Consequently any function $f(\mathcal{R})$ and any action of the $f(\mathcal{R})$ type will also be projective invariant. However, matter actions that depend on the connection will not be projective invariant. This has been discussed several times in the literature [5, 32, 33, 4, 10].

To resolve the inconsistency one needs to somehow break the projective invariance in the gravity sector. The only way to do that, given that we do not want to alter the form of the action, is to constraint the connection to some extent. The meaning of projective invariance is very similar to usual gauge invariance, in the sense that it implies that the connection can be determined only up to a projective transformation. So, to break gauge invariance we need a constraint that that would act as "gauge fixing". Clearly, given the nature of the projective transformation we essentially need to fix a vector. It has been argued in Refs. [4, 10] that the best choice for $f(\mathcal{R})$ gravity is to set

$$S_{\mu} \equiv S_{\alpha\mu}{}^{\alpha} = 0 \tag{39}$$

17

This constraint can be imposed implicitly, but also explicitly by adding to the action the Lagrange multiplier

$$S_{LM} = \int d^4x \sqrt{-g} B^\mu S_\mu \,. \tag{40}$$

Varying the total action with respect to metric $g^{\mu\nu}$, connection $\Gamma^\rho_{\ \mu\nu}$ and Lagrange Multiplier B^μ lead, after some simple manipulations [8, 4], to the following set of field equations

$$f'(\mathcal{R})\mathcal{R}_{(\mu\nu)} - \frac{1}{2}f(\mathcal{R})g_{\mu\nu} = \kappa T_{\mu\nu} \,, \tag{41}$$

$$-\nabla_\lambda(\sqrt{-g}f'(\mathcal{R})g^{\mu\nu}) + \nabla_\sigma\left(\sqrt{-g}f'(\mathcal{R})g^{\sigma\mu}\right)\delta^\nu_{\ \lambda} + 2\sqrt{-g}f'(\mathcal{R})(g^{\mu\nu}S_{\lambda\sigma}^{\ \ \sigma}$$
$$-g^{\mu\rho}\delta^\nu_{\ \lambda}S_{\rho\sigma}^{\ \ \sigma} + g^{\mu\sigma}S_{\sigma\lambda}^{\ \ \nu}) = \kappa\sqrt{-g}\left(\Delta_\lambda^{\ \mu\nu} - \frac{2}{3}\Delta_\sigma^{\ \sigma[\nu}\delta^{\mu]}_{\ \lambda}\right), \tag{42}$$

$$S_{\alpha\mu}^{\ \ \alpha} = 0 \,. \tag{43}$$

where a prime denotes differentiation with respect to the argument.

We now check whether it is possible to eliminate algebraically the connection from the field equations. This can be done following a similar procedure as the one used in section 3. A contraction of eq. (42) yields

$$\nabla_\sigma\sqrt{-g}f'(\mathcal{R})g^{\sigma\mu} = \kappa\frac{2}{3}\sqrt{-g}\Delta_\lambda^{\ (\mu\lambda)} \,. \tag{44}$$

We can use this equation in order to eliminate the second term in (42) to get

$$-\nabla_\lambda(\sqrt{-g}f'(\mathcal{R})g^{\mu\nu}) + 2\sqrt{-g}f'(\mathcal{R})g^{\mu\sigma}S_{\sigma\lambda}^{\ \ \nu} =$$
$$= \kappa\sqrt{-g}\left(\Delta_\lambda^{\ \mu\nu} - \frac{2}{3}\Delta_\sigma^{\ \sigma[\nu}\delta^{\mu]}_{\ \lambda} - \frac{2}{3}\Delta_\sigma^{\ (\mu\sigma)}\delta^\nu_{\ \lambda}\right) \,. \tag{45}$$

Using the identity

$$g_{\mu\nu}\partial_\lambda(\sqrt{-g}f'(\mathcal{R})g^{\mu\nu}) = 4\sqrt{-g}\partial_\lambda f'(\mathcal{R}) + 2f'(\mathcal{R})\sqrt{-g}\partial_\lambda \ln\sqrt{-g} \,, \tag{46}$$

and after contracting eq. (45) with the metric in the μ and ν indices one gets

$$\partial_\lambda \ln\sqrt{-g} = \frac{1}{2}\left[-\frac{\kappa}{f'}\left(g_{\mu\nu}\Delta_\lambda^{\ \mu\nu} - \frac{2}{3}g_{\mu\lambda}\Delta_\sigma^{\ (\mu\sigma)}\right) - 4\frac{\partial_\lambda f'}{f'} + 2\Gamma^\sigma_{\ \sigma\lambda}\right] \,. \tag{47}$$

18

Eliminating the density related terms and lowering the indices as we did in the previous section eq. (45) yields

$$-\partial_\lambda g_{\alpha\beta} - g_{\alpha\beta}\frac{\partial_\lambda f'}{f'} + \Gamma^\mu{}_{\beta\lambda}g_{\mu\alpha} + \Gamma^\mu{}_{\lambda\alpha}g_{\mu\beta} = \frac{\kappa}{f'}\left[\frac{1}{2}g_{\alpha\beta}g_{\mu\nu}\Delta_\lambda{}^{\mu\nu}\right. \tag{48}$$

$$\left. -\frac{1}{3}g_{\alpha\beta}g_{\mu\lambda}\Delta_\sigma{}^{(\mu\sigma)} - g_{\mu\alpha}g_{\nu\beta}\Delta_\lambda{}^{\mu\nu} + \frac{1}{3}\left(g_{\alpha\lambda}g_{\nu\beta}\Delta_\sigma{}^{\sigma\nu} + g_{\lambda\beta}g_{\mu\alpha}\Delta_\sigma{}^{\mu\sigma}\right)\right].$$

Adding suitable permutations of eq. (48) one gets

$$\Gamma^\rho{}_{\alpha\beta} = \left\{{}^\rho{}_{\alpha\beta}\right\} + \frac{1}{2f'}\left[\partial_\alpha f'\delta^\rho{}_\beta + \partial_\beta f'\delta^\rho{}_\alpha - g^{\rho\lambda}g_{\alpha\beta}\partial_\lambda f'\right] + \frac{\kappa}{f'}W_{\alpha\beta}{}^\rho, \tag{49}$$

where $\left\{{}^\rho{}_{\alpha\beta}\right\}$ is the usual Levi-Civita connection associated with the metric $g_{\mu\nu}$ and $W_{\alpha\beta}{}^\rho$ is a tensor encompassing all the hypermomenta terms

$$W_{\alpha\beta}{}^\rho = -\frac{1}{2}\left\{\frac{1}{2}g_{\alpha\beta}g_{\mu\nu}\Delta^{\rho\mu\nu} - g_{\mu\nu}\delta^\rho{}_{(\alpha}\Delta_{\beta)}{}^{\mu\nu} + \Delta_\beta{}^\rho{}_\alpha + \Delta_{\alpha\beta}{}^\rho\right. \tag{50}$$

$$-\Delta^\rho{}_{\alpha\beta} - g_{\alpha\beta}\Delta_\sigma{}^{(\rho\sigma)} + \frac{1}{3}\delta^\rho{}_\alpha g_{\mu\beta}\left(2\Delta_\sigma{}^{[\sigma\mu]} + \Delta_\sigma{}^{(\sigma\mu)}\right)$$

$$\left. +\frac{1}{3}\delta^\rho{}_\beta g_{\mu\alpha}\left(2\Delta_\sigma{}^{[\sigma\mu]} + \Delta_\sigma{}^{(\sigma\mu)}\right)\right\}.$$

Eq. (49) provides and expression for the connection in terms of the metric, the hypermomentum but also $\mathcal{R}$, via the presence of f. So, we essentially run into the same difficulties we faced in the previous section when trying to eliminate the connection. However, here $\mathcal{R}$ is just a scalar quantity. Consider the trace of eq. (41)

$$\mathcal{R}f'(\mathcal{R}) - 2f(\mathcal{R}) = \kappa T. \tag{51}$$

For a given function f this is an algebraic equation in $\mathcal{R}$. Setting aside pathological cases in which this equation has no root, and the exceptional case where $f(\mathcal{R}) \propto \mathcal{R}^2$, which corresponds to a conformally invariant gravitational action (see Refs. [34, 26, 4] for more details), eq. (51) can be used to express $\mathcal{R}$ as an algebraic function of T. This expression can in turn be used to eliminate $\mathcal{R}$ in favour of T in the f terms in eq. (49). Therefore, from now on we can be thinking of eq. (49) as expressing the affine connection as a function of just derivatives of metric, $T_{\mu\nu}$ and the hypermomentum. This mean we are clear of the first difficulty encountered for generic fourth order actions.

This is not the case for the second point we made previously though, *i.e.* that $T_{\mu\nu}$, and hence T, can generically depend on the connection. Even though the requirement that the matter satisfies second order differential equations of motion essentially implies that the dependence of $T_{\mu\nu}$ on the connection will be algebraic (there can be only first covariant derivatives of the matter fields is S_M), the fact that there are first order derivatives of f' in eq. (49) is enough to give derivatives of the connection. Therefore, in metric-affine $f(\mathcal{R})$ the connection satisfies a dynamical equation in general.

A remarkable observation is the following. Taking the antisymmetric part of (49) in its lower indices we get

$$\Gamma^{\rho}_{[\alpha\beta]} \equiv S_{\alpha\beta}{}^{\rho} = \Delta_{[\beta}{}^{\rho}{}_{\alpha]} + \Delta_{[\alpha\beta]}{}^{\rho} - \Delta^{\rho}{}_{[\alpha\beta]} \tag{52}$$
$$= g^{\rho\lambda}\left(\Delta_{\beta[\lambda\alpha]} + \Delta_{\alpha[\beta\lambda]} - \Delta_{\lambda[\alpha\beta]}\right).$$

This implies that the torsion is still non-dynamical and vanishes for matter fields with vanishing $\Delta_{\gamma}{}^{[\alpha\beta]}$. It is only the symmetric part of the connection that carries dynamics. As already stressed in [4] torsion is non-propagating in metric-affine $f(\mathcal{R})$ and it is introduced by matter fields having $\Delta_{\gamma}{}^{[\alpha\beta]} \neq 0$.

The fact that the connection appears to satisfy a first order differential equation, at least if one assumes that T does not include any derivatives of the connection, seems worrying. However, it is very difficult to tell if this is indeed a problem. Neither do we have the exact form of the equation, nor do we know which degrees of freedom hiding in the connection will actually be excited.

Of course, for matter fields which do not couple to the connection (scalar field, gauge fields) or if one imposes that the independent connection does not enter the matter action S_M, $T_{\mu\nu}$ is independent of the connection as well and $\Delta_{\lambda}{}^{\mu\nu} = 0$. In this case the connection can indeed be eliminated and one recovers the results of Palatini $f(\mathcal{R})$ gravity [25]. Another special case is the one where $f(\mathcal{R}) = \mathcal{R}$, as is this case $f' = 1$ and $\mathcal{R}$ is no longer present in the eq. (49), which now takes the form

$$\Gamma^{\rho}_{\alpha\beta} = \left\{{}^{\rho}_{\alpha\beta}\right\} + \kappa\, W_{\alpha\beta}{}^{\rho} . \tag{53}$$

One can then write

$$\mathcal{R}_{(\alpha\beta)} \equiv \partial_{\rho}\Gamma^{\rho}_{(\alpha\beta)} - \partial_{(\beta}\Gamma^{\rho}_{\alpha)\rho} + \Gamma^{\rho}_{\sigma\rho}\Gamma^{\sigma}_{(\alpha\beta)} - \Gamma^{\rho}_{\sigma(\beta}\Gamma^{\sigma}_{\alpha)\rho} = \tag{54}$$
$$= R_{\alpha\beta} + \kappa\left[\bar{\nabla}_{\rho}W_{(\alpha\beta)}{}^{\rho} - \bar{\nabla}_{(\beta}W_{\alpha)\rho}{}^{\rho} + W_{\sigma\rho}{}^{\rho}W_{(\alpha\beta)}{}^{\sigma} - W_{\sigma(\beta}{}^{\rho}W_{\alpha)\rho}{}^{\sigma}\right],$$

20

where $R_{\mu\nu}$ is the Ricci tensor of the metric $g_{\mu\nu}$ and $\bar{\nabla}_\mu$ is the covariant derivative defined with the Levi-Civita connection of the same metric. Contracting with the metric one gets

$$\mathcal{R} \;=\; R + \kappa\left[2\bar{\nabla}_{[\rho}W^{\mu}{}_{\mu]}{}^{\rho} + W_{\sigma\rho}{}^{\rho}W_\mu{}^{\mu\sigma} - W_\sigma{}^{\mu\rho}W_{\mu\rho}{}^{\sigma}\right]. \tag{55}$$

We can now use eqs. (53), (54) and (55) in order to completely eliminate the connection and end up with the single field equation for the metric

$$\begin{aligned}
G_{\alpha\beta} \;&=\; \kappa\, T_{\alpha\beta} + \frac{\kappa}{2}g_{\alpha\beta}\left\{2\bar{\nabla}_{[\rho}W^{\mu}{}_{\mu]}{}^{\rho} + W_{\sigma\rho}{}^{\rho}W_\mu{}^{\mu\sigma} - W_\sigma{}^{\mu\rho}W_{\mu\rho}{}^{\sigma}\right\} \\
&\quad - \kappa\left\{\bar{\nabla}_\rho W_{(\alpha\beta)}{}^{\rho} - \bar{\nabla}_{(\beta}W_{\alpha)\rho}{}^{\rho} + W_{\sigma\rho}{}^{\rho}W_{(\alpha\beta)}{}^{\sigma} - W_{\sigma(\beta}{}^{\rho}W_{\alpha)\rho}{}^{\sigma}\right\}, \tag{56}
\end{aligned}$$

where, as usual,

$$G_{\alpha\beta} \equiv R_{\alpha\beta} - \frac{1}{2}R\,g_{\alpha\beta}\,, \tag{57}$$

is the Einstein tensor of the metric $g_{\mu\nu}$. Therefore, $f(\mathcal{R}) = \mathcal{R}$ metric-affine gravity reduces to general relativity with extra matter interactions. This is anyway clearly just a subcase of the most general second order action we examined in section 3 with vanishing a_i's.

However, we have shown for any other function $f(\mathcal{R})$ the connection cannot be algebraically eliminated in the presence of matter fields that couple to it.

6. Discussion and Conclusions

We have studied the dynamics of theories of gravity in which the metric and the connection are independent quantities. Instead of restricting ourselves to a specific action, which would inevitably affect the generality of our conclusions, we chose to follow an approach inspired by effective field theory and attempt to understand how are the dynamics of the theory affected when increasing the order of the various invariant included in the action. To this end we first considered the most general action formed by second order invariants and then moved on to examine how these would be modified by including different types of higher order terms in the action. In both cases we imposed a generalized minimal coupling principle in order to reduce the number of terms to be considered, which excludes invariants constructed with the non-metricity or the metric curvature.

Our main conclusions are the following:

21

- Even for the most general action one can construct with second order invariants the connection does not carry any dynamics and can always be algebraically eliminated. That is, at this order, metric-affine gravity can always be written as general relativity with a modified source term or extra matter interactions. No extra degrees of freedom are excited.

- Including higher order terms in the action changes the situation radically. The connection (or parts of it) becomes dynamical and so, it cannot be eliminated algebraically. The theory now propagates more degrees of freedom than general relativity. Thus, seen as an effective field theory, metric-affine gravity is rather peculiar and its dynamics can deceive: at the lowest order the extra degrees of freedom appear to lose their dynamics and become auxiliary fields, but once higher order terms are taken into account the extra degrees of freedom do propagate. To avoid exciting extra degrees of freedom significant fine tuning and extra *a priori* constraints are required.

- $f(\mathcal{R})$ actions, which have been previously considered in metric-affine gravity, appear to constitute a distinct class with special properties. Even though the connection does carry dynamics in the presence of fields coupling to it — unlike the simplified case of Palatini $f(\mathcal{R})$ gravity — torsion remains non-propagating. The propagating degrees of freedom reside only in the symmetric part of the connection. In this sense, $f(\mathcal{R})$ actions cannot be considered representative examples of generic higher order metric-affine theories.

From an effective field theory perspective it seems that there are dynamical degrees of freedom in metric-affine gravity which appear to "freeze" at low energies and can be eliminated in favour of extra matter interaction. This implies that a possible low energy manifestation of metric-affine gravity could be revealed in matter experiments in terms of such interactions, but the phenomenology of metric-affine theories is not limited to that. It is much richer and it includes extra propagating degrees of freedom, which can potentially be detected. A typical, but certainly not the only, example would be the presence of propagating torsion, whose consequences have been studied in a limiting setting [35].[3]

---

[3]See also Ref. [36], which appeared during the completion of this manuscript, and references therein.

It goes beyond the scope of the current study to examine further the phenomenology of metric-affine gravity. It would be very interesting to understand in more detail how the extra degrees of freedom behave in the regime where they are dynamical and how exactly do they modify matter interactions at low energies. It is also crucial to examine the predictions of such theories for energy conservation and violations of the various formulations of the equivalence principle. Such considerations would allow us to place constraints on metric-affine theories.

A probable point of concern can be our use of the generalized minimal coupling principle. One could argue that it is not compatible with our effective field theory perspective as radiative corrections would not respect such a principle. One could also feel uneasy treating non-metricity and torsion on a different footing. Indeed, the minimal coupling principle is used here mostly as a way to reduce the number of terms one has to take into consideration and it should not necessarily be considered as a fundamental principle. Abandoning it and considering the most general action possible is the next step.

As a closing remark, we would like to mention the following. Clearly, one might question how fundamental is the geometrical interpretation of metric-affine gravity. In fact, since for second order actions one can always eliminate the independent connection, the latter can be regarded as an auxiliary field. Even for actions with higher order terms though, where degrees of freedom residing in the connection will be excited, one could have an equivalent representation without an independent affine connection (recall that an independent connection can always be written as the Levi-Civita connection plus a tensor). Indeed, which representation one choose is a matter of preference, at least at a classical level, as the dynamical content of the theory is one and the same. On the other hand, it is worth pointing out that the choice of representations becomes a factor when constructing the action of the theory. It influences our choices regarding the presence of some terms by making some exclusion principles, such as minimal coupling and its generalizations, more or less appealing (see also Ref. [24] for a more general discussion on this issue). This is a subtle point that needs to be taken seriously into account.

Acknowledgments

for discussions. TPS was supported in part by the STFC and in part by a Marie Curie Incoming International Fellowship.

References

[1] C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation*, (Freeman and Co., San Francisco, 1973).

[2] R. M. Wald, *General Relativity*, (The University of Chicago Press, Chicago, 1984).

[3] T. P. Sotiriou, *Proceedings of the Eleventh Marcel Grossmann Meeting on General Relativity*, edited by H. Kleinert, R.T. Jantzen and R. Ruffini, p. 1223-1226, (World Scientific, Singapore, 2008), arXiv: gr-qc/0611158.

[4] T. P. Sotiriou and S. Liberati, Annals Phys. **322** (2007) 935 [arXiv:gr-qc/0604006]; J. Phys. Conf. Ser. **68** (2007) 012022. [arXiv:gr-qc/0611040].

[5] F. W. Hehl and G. D. Kerling, Gen. Rel. Grav. **9** (1978) 691.

[6] H. A. Buchdahl, Mon. Not. Roy. Ast. Soc. **150** (1970) 1.

[7] D. Barraco, V. H. Hamity and H. Vucetich, Gen. Rel. Grav. **34** (2002) 533; D. N. Vollick, Phys. Rev. D **68** (2003) 063510; G. Allemandi, A. Borowiec and M. Francaviglia, Phys. Rev. D **70** (2004) 043524; G. Allemandi, A. Borowiec, M Francaviglia and S. D. Odintsov, Phys. Rev. D **72** (2005) 063505; T. P. Sotiriou, Phys. Rev. D **73** (2006) 063515; M. Amarzguioui, O. Elgaroy, D. F. Mota and T. Multamaki, Astron. Astrophys. **454** (2006) 707; T. P. Sotiriou, Class. Quant. Grav. **23** (2006) 1253; T. Koivisto and H. Kurki-Suonio, Class. Quant. Grav. **23** (2006) 2355; T. Koivisto, Phys. Rev. D **73** (2006) 083517.

[8] T. P. Sotiriou and V. Faraoni, arXiv:0805.1726 [gr-qc].

[9] A. De Felice and S. Tsujikawa, arXiv:1002.4928 [gr-qc].

[10] T. P. Sotiriou, *Modified Actions for gravity: Theory and Phenomenology*, PhD thesis (SISSA, Trieste, 2007) arXiv: 0710.4438 [gr-qc].

[11] E. E. Flanagan, Phys. Rev. Lett. **92**, 071101 (2004); Class. Quant. Grav. **21** (2004) 3817.

[12] A. Iglesias, N. Kaloper, A. Padilla and M. Park, Phys. Rev. D **76** (2007) 104001.

[13] G. J. Olmo, Phys. Rev. Lett. **95** (2005) 261102.

[14] T. P. Sotiriou, Gen. Rel. Grav. **38** (2006) 1407.

[15] E. Barausse, T. P. Sotiriou and J. C. Miller, Class. Quant. Grav. **25** (2008) 062001; Class. Quant. Grav. **25** (2008) 105008; EAS Publ. Ser. **30** (2008) 189.

[16] G. J. Olmo, Phys. Rev. D **78** (2008) 104026.

[17] G. Allemandi, A. Borowiec and M. Francaviglia, Phys. Rev. D **70** (2004) 103503.

[18] B. Li, J. D. Barrow and D. F. Mota, Phys. Rev. D **76** (2007) 104047.

[19] G. J. Olmo, H. Sanchis-Alepuz and S. Tripathi, Phys. Rev. D **80** (2009) 024013.

[20] C. Barragan and G. J. Olmo, arXiv:1005.4136v1 [gr-qc].

[21] F. Bauer, arXiv:1007.2546v1 [gr-qc].

[22] V. Vitagliano, T. P. Sotiriou and S. Liberati, arXiv:1007.3937 [gr-qc].

[23] T. P. Sotiriou, Phys. Lett. B **645** (2007) 389.

[24] T. P. Sotiriou, V. Faraoni and S. Liberati, *Int. J. Mod. Phys. D* **17** (2008) 399.

[25] T. P. Sotiriou, Class. Quant. Grav. **26** (2009) 152001.

[26] T. P. Sotiriou, Class. Quant. Grav. **23** (2006) 5117.

[27] F. W. Hehl, P. Von Der Heyde, G. D. Kerlick and J. M. Nester, Rev. Mod. Phys. **48** (1976) 393.

[28] Yu. N. Obukhov, E. J. Vlachynsky, W. Esser, and F. W. Hehl, Phys. Rev. D **56** (1997) 7769.

[29] S. M. Christensen, J. Phys. A **13** (1980) 3001.

[30] A. Papapetrou and J. Stachel, Gen. Rel. Grav. **9** (1978) 1075.

[31] F. W. Hehl, G. D. Kerlick and P. Von Der Heyde, Z. Naturforsch. **31A** (1976) 524

[32] E. Schrödinger, *Space-Time Structure*, (Cambridge University Press, Cambridge, 1963).

[33] V. D. Sandberg, Phys. Rev. D **12** (1975) 3013.

[34] M. Ferraris, M. Francaviglia and I. Volovich, arXiv: gr-qc/9303007.

[35] S. M. Carroll and G. B. Field, Phys. Rev. D **50** (1994) 3867.

[36] I. L. Shapiro, arXiv:1007.5294 [hep-th].